

SEQUENCE LISTING

5 (1) GENERAL INFORMATION:

- 10 (i) APPLICANTS: Kendall, Richard L.
Thomas, Kenneth A.
Mao, Xianzhi
Tebben, Andrew J.
- (ii) TITLE OF INVENTION: HUMAN RECEPTOR TYROSINE KINASE, KDR
- 15 (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
20 (A) ADDRESSEE: Merck & Co., Inc.
(B) STREET: P.O. Box 2000
(C) CITY: Rahway
(D) STATE: NJ
(E) COUNTRY: US
(F) ZIP: 07065-0907
- 25 (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- 30 (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
- 35 (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Hand, J. Mark
(B) REGISTRATION NUMBER: 36,545
(C) REFERENCE/DOCKET NUMBER: 19963PV
- 40 (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 732/594-3905
(B) TELEFAX: 732/594-4720
- 45 (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:
50 (A) LENGTH: 4071 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

	ATGGAGAGCA AGGTGCTGCT GGCCGTCGCC CTGTGGCTCT GCGTGGAGAC CCGGGCCGCC	60
5	TCTGTGGGTT TGCCTAGTGT TTCTCTTGAT CTGCCCCAGC TCAGCATACA AAAAGACATA	120
	CTTACAATTA AGGCTAATAC AACTCTTCAA ATTACTTGCA GGGGACAGAG GGACTTGGAC	180
10	TGGCTTTGGC CCAATAATCA GAGTGGCAGT GAGCAAGGG TGGAGGTGAC TGAGTGCAGC	240
	GATGGCTCTT TCTGTAAGAC ACTCACAATT CCAAAAGTGA TCGGAAATGA CACTGGAGCC	300
	TACAAGTGCT TCTACCGGGA AACTGACTTG GCCTCGGTCA TTTATGTCTA TGTTCAAGAT	360
15	TACAGATCTC CATTATTGTC TTCTGTTAGT GACCAACATG GAGTCGTGTA CATTACTGAG	420
	AACAAAAACA AAAGTGTGGT GATTCATGTT CTCGGGTCCA TTTCAAATCT CAACGTGTCA	480
20	CTTTGTGCAA GATACCCAGA AAAGAGATTT GTTCCTGATG GTAACAGAAT TTCCTGGGAC	540
	AGCAAGAAGG GCTTTACTAT TCCCAGCTAC ATGATCAGCT ATGCTGGCAT GGTCTTCTGT	600
	GAAGCAAAAA TTAATGATGA AAGTTACCAG TCTATTATGT ACATAGTTGT CGTTGTAGGG	660
25	TATAGGATTT ATGATGTGGT TCTGAGTCCG TCTCATGGAA TTGAACATAT TGTTGGAGAA	720
	AAGCTTGTCT TAAATTGTAC AGCAAGAAGT GAACATAAAT TGGGGATTGA CTTCAACTGG	780
30	GAATACCCCT CTTCGAAGCA TCAGCATAAG AAACCTTGTA ACCGAGACCT AAAAACCCAG	840
	TCTGGGAGTG AGATGAAGAA ATTTTGTGAG ACCTTAACTA TAGATGTTGT AATCCGGAGT	900
	GACCAAGGAT GTTACACCTG TGCAGCATCC AGTGGGCTGA TGACCAAGAA GAACAGCACA	960
35	TTTGTCAAGG TCCATGAAAA ACCTTTTGTG GCTTTTGGAA GTGGCATGGA ATCTCTGGTG	1020
	GAAGCCACGG TGGGGAGCG TGTGAGAATC CTTGCGAAGT ACCTTGGTTA CCCACCCCA	1080
40	GAAATAAAAT GGTATAAAAA TGGAATACCC CTGAGTCCA ATCACACAAT TAAAGCGGG	1140
	CATGTACTGA CGATTATGGA AGTGAGTGAA AGAGACACAG GAAATTACAC TGTATCCTT	1200
	ACCAATCCCA TTTCAAAGGA GAAGCAGAGC CATGTGGTCT CTCTGGTGTG GTATGTCCCA	1260
45	CCCCAGATTG GTGAGAAATC TCTAATCTCT CCTGTGGATT CCTACCAGTA CGGCACCACT	1320
	CAAACGCTGA CATGTACGGT CTATGCCATT CCTCCCCGCG ATCACATCCA CTGGTATTGG	1380
50	CAGTTGGAGG AAGAGTGCGC CAACGAGCCC AGCCAAGCTG TCTCAGTGAC AAACCCATAC	1440
	CCTTGTGAAG AATGGAGAAG TGTGGAGGAC TTCCAGGGAG GAAATAAAAT TGAAGTTAAT	1500
	AAAAATCAAT TTGCTCTAAT TGAAGGAAAA AACAAAAGTG TAAGTACCCT TGTATTCCAA	1560
55	GCGGCAAAAT GTGACGTTT GTACAAATGT GAAGCGGTCA ACAAGATCGG GAGAGGAGAG	1620
	AGGGTGATCT CCTTCCACGT GACCAAGGGT CCGTGAATTA CTTTGCAACC TGACATGCAG	1680

	CCCACTGAGC	AGGAGAGCGT	GTCCTTGTGG	TGCACTGCAG	ACAGATCTAC	GTTTGAGAAC	1740
	CTCACATGST	ACAAGCTTGG	CCCACAGCCT	CTGCCAATCC	ATGTGGGAGA	GTTGCCCCACA	1800
5	CCTGTTTGCA	AGAACTTGGA	TACTCTTTGG	AAATTGAATG	CCACCATGTT	CTCTAATAGC	1860
	ACAAATGACA	TTTGTATCAT	GGAGCTTAAG	AATGCATCCT	TGCAGGACCA	AGGAGACTAT	1920
10	GTCTGCCTTG	CTCAAGACAG	GAAGACCAAG	AAAAGACATT	GCGTGGTCAG	GCAGCTCACA	1980
	GTCTTAGAGC	GTGTGGCACC	CACGATCACA	GGAAACCTGG	AGAATCAGAC	GACAAGTATT	2040
	GGGGAAGCA	TCGAAGTCTC	ATGCACGGCA	TCTGGGAATC	CCCCCTCCACA	GATCATGTGG	2100
15	TTTAAAGATA	ATGAGACCCCT	TGTAGAAGAC	TCAGGCATTG	TATTGAAGGA	TGGGAACCGG	2160
	AACCTCACTA	TCCGCAGAGT	GAGGAAGGAG	GACGAAGGCC	TCTACACCTG	CCAGGCATGC	2220
20	AGTGTTCCTG	GCTGTGCAAA	AGTGGAGGCA	TTTTTTCATA	TAGAAGGTGC	CCAGGAAAAG	2280
	ACGAACTTGG	AAATCATTAT	TCTAGTAGGC	ACGGCGGTGA	TTGCCATGTT	CTTCTGGCTA	2340
	CTTCTTGTC	TCATCCTACG	GACCGTTAAG	CGGGCCCAATG	GAGGGGAAC	GAAGACAGGG	2400
25	TACCTGTCCA	TCGTCATGGA	TCCAGATGAA	CTCCCATTGG	ATGAACATTG	TGAACGACTG	2460
	CCTTATGATG	CCAGCAAAAT	GGAATTCCCC	AGAGACCGGC	TGAAGCTAGG	TAAGCCTCTT	2520
30	GGCCGTGGTG	CCTTTGGCCA	AGTGATTGAA	GCAGATGCCT	TTGGAATTGA	CAAGACAGCA	2580
	ACTTGCAGGA	CAGTAGCAGT	CAAAATGTTG	AAAGAAGGAG	CAACACACAG	TGAGCATCGA	2640
	GCTCTCATGT	CTGAACCTCA	GATCCTCATT	CATATTGGTC	ACCATCTCAA	TGTGGTCAAC	2700
35	CTTCTAGGTG	CCTGTACCAA	GCCAGGAGGG	CCACTCATGG	TGATTGTGGA	ATTCTGCAAA	2760
	TTTGAAACC	TGTCCACTTA	CCTGAGGAGC	AAGAGAAATG	AATTTGTCCC	CTACAAGACC	2820
40	AAAGGGGCAC	GATTCCGTCA	AGGGAAGAC	TACGTTGGAG	CAATCCCTGT	GGATCTGAAA	2880
	CGCGCTTGG	ACAGCATCAC	CAGTAGCCAG	AGCTCAGCCA	GCTCTGGATT	TGTGGAGGAG	2940
	AAGTCCCTCA	GTGATGTAGA	AGAAGAGGAA	GCTCCTGAAG	ATCTGTATAA	GGACTTCCTG	3000
45	ACCTTGGAGC	ATCTCATCTG	TTACAGCTTC	CAAGTGGCTA	AGGCATGGA	GTTCTTGGCA	3060
	TCGCGAAAGT	GTATCCACAG	GGACCTGGCG	GCACGAAATA	TCCTCTTATC	GGAGAAGAAC	3120
50	GTGGMATAAA	TCTGTGACTT	TGGCTTGGCC	CGGATATTTT	ATAAAGATCC	AGATTATGTC	3180
	AGAAAAGGAG	ATGCTCGCCT	CCCTTTGAAA	TGGATGGCCC	CAGAAACAAT	TTTTTGACAGA	3240
	GTGTACACAA	TCCAGAGTGA	CGTCTGGTCT	TTTGGTGTGT	TGCTGTGGGA	AATATTTTCC	3300
55	TTAGGTGCTT	CTCCATATCC	TGGGGTAAAG	ATTGATGAAG	AATTTTGTAG	GCGATTGAAA	3360
	GAAGGAAC	GAATGAGGGC	CCCTGATTAT	ACTACACCA	AAATGTACCA	GACCATGCTG	3420

GACTGCTGGC ACGGGGAGCC CAGTCAGAGA CCCACGTTT CAGAGTTGGT GGAACATTG 3480
 GGAAATCTCT TCGAAGCTAA TGCTCAGCAG GATGGCAAAG ACTACATTGT TCTTCCGATA 3540
 5 TCAGAGACTT TGAGCATGGA AGAGGATTCT GGACTCTCTC TGCTACCTC ACCTGTTTCC 3600
 TGTATGGAGG AGGAGGAAGT ATGTGACCCC AAATTCCATT ATGACAACAC AGCAGGAATC 3660
 AGTCAGTATC TGCAGAACAG TAAGCGAAAG AGCCGGCCTG TGAGTGTAAG AACATTGTA 3720
 10 GATATCCCGT TAGAAGAACC AGAAGTAAAA GTAATCCCAG ATGACAACCA GACGGACAGT 3780
 GGTATGGTTC TTGCTCAGA AGAGCTGAAA ACTTTGGAAG ACAGAACCAA ATTATCTCCA 3840
 15 TCITTTGGTG GAATGGTGCC CAGCAAAAGC AGGGAGTCTG TGGCATCTGA AGGCTCAAAC 3900
 CAGACAAGCG GCTACCAGTC CGGATATCAC TCCGATGACA CAGACACCAC CGTGTACTCC 3960
 AGTGAGGAAG CAGAACTTTT AAAGCTGATA GAGATTGGAG TGCAAACCGG TAGCACAGCC 4020
 20 CAGATTCTCC AGCCTGACTC GGGGACCACA CTGAGCTCTC CTCTGTTTA A 4071

(2) INFORMATION FOR SEQ ID NO:2:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1356 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
 Met Glu Ser Lys Val Leu Leu Ala Val Ala Leu Trp Leu Cys Val Glu
 1 5 10 15
 35 Thr Arg Ala Ala Ser Val Gly Leu Pro Ser Val Ser Leu Asp Leu Pro
 20 25 30
 40 Arg Leu Ser Ile Gln Lys Asp Ile Leu Thr Ile Lys Ala Asn Thr Thr
 35 40 45
 Leu Gln Ile Thr Cys Arg Gly Gln Arg Asp Leu Asp Trp Leu Trp Pro
 50 55 60
 45 Asn Asn Gln Ser Gly Ser Glu Gln Arg Val Glu Val Thr Glu Cys Ser
 65 70 75 80
 Asp Gly Leu Phe Cys Lys Thr Leu Thr Ile Pro Lys Val Ile Gly Asn
 85 90 95
 50 Asp Thr Gly Ala Tyr Lys Cys Phe Tyr Arg Glu Thr Asp Leu Ala Ser
 100 105 110
 Val Ile Tyr Val Tyr Val Gln Asp Tyr Arg Ser Pro Phe Ile Ala Ser
 115 120 125
 55 Val Ser Asp Gln His Gly Val Val Tyr Ile Thr Glu Asn Lys Asn Lys
 130 135 140

Thr Val Val Ile Pro Cys Leu Gly Ser Ile Ser Asn Leu Asn Val Ser
 145 150 155 160
 5 Leu Cys Ala Arg Tyr Pro Glu Lys Arg Phe Val Pro Asp Gly Asn Arg
 165 170 175
 Ile Ser Trp Asp Ser Lys Lys Gly Phe Thr Ile Pro Ser Tyr Met Ile
 180 185 190
 10 Ser Tyr Ala Gly Met Val Phe Cys Glu Ala Lys Ile Asn Asp Glu Ser
 195 200 205
 Tyr Gln Ser Ile Met Tyr Ile Val Val Val Val Gly Tyr Arg Ile Tyr
 210 215 220
 15 Asp Val Val Leu Ser Pro Ser His Gly Ile Glu Ser Val Gly Glu
 225 230 235 240
 20 Lys Leu Val Leu Asn Cys Thr Ala Arg Thr Glu Leu Asn Val Gly Ile
 245 250 255
 Asp Phe Asn Trp Glu Tyr Pro Ser Ser Lys His Gln His Lys Lys Leu
 260 265 270
 25 Val Asn Arg Asp Leu Lys Thr Gln Ser Gly Ser Glu Met Lys Lys Phe
 275 280 285
 30 Leu Ser Thr Leu Thr Ile Asp Gly Val Thr Arg Ser Asp Gln Gly Leu
 290 295 300
 Tyr Thr Cys Ala Ala Ser Ser Gly Leu Met Thr Lys Lys Asn Ser Thr
 305 310 315 320
 35 Phe Val Arg Val His Glu Lys Pro Phe Val Ala Phe Gly Ser Gly Met
 325 330 335
 Glu Ser Leu Val Glu Ala Thr Val Gly Glu Arg Val Arg Ile Pro Ala
 340 345 350
 40 Lys Tyr Leu Gly Tyr Pro Pro Pro Glu Ile Lys Trp Tyr Lys Asn Gly
 355 360 365
 45 Ile Pro Leu Glu Ser Asn His Thr Ile Lys Ala Gly His Val Leu Thr
 370 375 380
 Ile Met Glu Val Ser Glu Arg Asp Thr Gly Asn Tyr Thr Val Ile Leu
 385 390 395 400
 50 Thr Asn Pro Ile Ser Lys Glu Lys Gln Ser His Val Val Ser Leu Val
 405 410 415
 Val Tyr Val Pro Pro Gln Ile Gly Glu Lys Ser Leu Ile Ser Pro Val
 420 425 430
 55 Asp Ser Tyr Gln Tyr Gly Thr Thr Gln Thr Leu Thr Cys Thr Val Tyr
 435 440 445

Ala Ile Pro Pro Pro His His Ile His Trp Tyr Trp Gln Leu Glu Glu
 450 455 460
 5 Glu Cys Ala Asn Glu Pro Ser Gln Ala Val Ser Val Thr Asn Pro Tyr
 465 470 475 480
 Pro Cys Glu Glu Trp Arg Ser Val Glu Asp Phe Gln Gly Gly Asn Lys
 485 490 495
 10 Ile Glu Val Asn Lys Asn Gln Phe Ala Leu Ile Glu Gly Lys Asn Lys
 500 505 510
 Thr Val Ser Thr Leu Val Ile Gln Ala Ala Asn Val Ser Ala Leu Tyr
 515 520 525
 15 Lys Cys Glu Ala Val Asn Lys Val Gly Arg Gly Glu Arg Val Ile Ser
 530 535 540
 Phe His Val Thr Arg Gly Pro Glu Ile Thr Leu Gln Pro Asp Met Gln
 545 550 555 560
 20 Pro Thr Glu Gln Glu Ser Val Ser Leu Trp Cys Thr Ala Asp Arg Ser
 565 570 575
 25 Thr Phe Glu Asn Leu Thr Trp Tyr Lys Leu Gly Pro Gln Pro Leu Pro
 580 585 590
 Ile His Val Gly Glu Leu Pro Thr Pro Val Cys Lys Asn Leu Asp Thr
 595 600 605
 30 Leu Trp Lys Leu Asn Ala Thr Met Phe Ser Asn Ser Thr Asn Asp Ile
 610 615 620
 Leu Ile Met Glu Leu Lys Asn Ala Ser Leu Gln Asp Gln Gly Asp Tyr
 625 630 635 640
 35 Val Cys Leu Ala Gln Asp Arg Lys Thr Lys Lys Arg His Cys Val Val
 645 650 655
 40 Arg Gln Leu Thr Val Leu Glu Arg Val Ala Pro Thr Ile Thr Gly Asn
 660 665 670
 Leu Glu Asn Gln Thr Thr Ser Ile Gly Glu Ser Ile Glu Val Ser Cys
 675 680 685
 45 Thr Ala Ser Gly Asn Pro Pro Pro Gln Ile Met Trp Phe Lys Asp Asn
 690 695 700
 Glu Thr Leu Val Glu Asp Ser Gly Ile Val Leu Lys Asp Gly Asn Arg
 705 710 715 720
 50 Asn Leu Thr Ile Arg Arg Val Arg Lys Glu Asp Glu Gly Leu Tyr Thr
 725 730 735
 55 Cys Gln Ala Cys Ser Val Leu Gly Cys Ala Lys Val Glu Ala Phe Phe
 740 745 750

5 Ile Ile Glu Gly Ala Gln Glu Lys Thr Asn Leu Glu Ile Ile Ile Leu
 755 760 765
 Val Gly Thr Ala Val Ile Ala Met Phe Phe Trp Leu Leu Leu Val Ile
 770 775 780
 10 Ile Leu Arg Thr Val Lys Arg Ala Asn Gly Gly Glu Leu Lys Thr Gly
 785 790 795 800
 Tyr Leu Ser Ile Val Met Asp Pro Asp Glu Leu Pro Leu Asp Glu His
 805 810 815
 15 Cys Glu Arg Leu Pro Tyr Asp Ala Ser Lys Trp Glu Phe Pro Arg Asp
 820 825 830
 Arg Leu Lys Leu Gly Lys Pro Leu Gly Arg Gly Ala Phe Gly Gln Val
 835 840 845
 20 Ile Glu Ala Asp Ala Phe Gly Ile Asp Lys Thr Ala Thr Cys Arg Thr
 850 855 860
 Val Ala Val Lys Met Leu Lys Glu Gly Ala Thr His Ser Glu His Arg
 865 870 875 880
 25 Ala Leu Met Ser Glu Leu Lys Ile Leu Ile His Ile Gly His His Leu
 885 890 895
 30 Asn Val Val Asn Leu Leu Gly Ala Cys Thr Lys Pro Gly Gly Pro Leu
 900 905 910
 Met Val Ile Val Glu Phe Cys Lys Phe Gly Asn Leu Ser Thr Tyr Leu
 915 920 925
 35 Arg Ser Lys Arg Asn Glu Phe Val Pro Tyr Lys Thr Lys Gly Ala Arg
 930 935 940
 Phe Arg Gln Gly Lys Asp Tyr Val Gly Ala Ile Pro Val Asp Leu Lys
 945 950 955 960
 40 Arg Arg Leu Asp Ser Ile Thr Ser Ser Gln Ser Ser Ala Ser Ser Gly
 965 970 975
 Phe Val Glu Glu Lys Ser Leu Ser Asp Val Glu Glu Glu Glu Ala Pro
 980 985 990
 45 Glu Asp Leu Tyr Lys Asp Phe Leu Thr Leu Glu His Leu Ile Cys Tyr
 995 1000 1005
 50 Ser Phe Gln Val Ala Lys Gly Met Glu Phe Leu Ala Ser Arg Lys Cys
 1010 1015 1020
 Ile His Arg Asp Leu Ala Ala Arg Asn Ile Leu Leu Ser Glu Lys Asn
 1025 1030 1035 1040
 55 Val Val Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile Tyr Lys Asp
 1045 1050 1055

- 46 -

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGAATTCCAT CCAAGCGGCA AATGTGTC

28

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGAATTCGGA GTCTTCTACA AGGGTCTC

28

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TTATGACAAC ACAGCAGG

18

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

5 TTGGATCCTC GAGTTGGGGT GTGGATGC 28

(2) INFORMATION FOR SEQ ID NO:7:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

20 GGATCCAGAT GAACTCCCAT TG 22

(2) INFORMATION FOR SEQ ID NO:8:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

35 GTCGACTTAG TCTTTGCCAT CCTGCTGAGC 30

40